

#4

TECH CENTER 1600/2900

DEC 18 2001

RECEIVED

RAW SEQUENCE LISTING

DATE: 11/29/2001

PATENT APPLICATION: US/09/938,885A

TIME: 11:02:26

Input Set : N:\Crf3\RULE60\09938885A.RAW

Output Set: N:\CRF3\11292001\I938885A.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Bandman, Olga

6 Goli, Surya K.

7 Murry, Lynn E.

9 (ii) TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT

11 (iii) NUMBER OF SEQUENCES: 5

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

15 (B) STREET: 3174 Porter Drive

16 (C) CITY: Palo Alto

17 (D) STATE: CA

18 (E) COUNTRY: US

19 (F) ZIP: 94304

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Diskette

23 (B) COMPUTER: IBM Compatible

24 (C) OPERATING SYSTEM: DOS

25 (D) SOFTWARE: FastSEQ Version 2.0

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/09/938,885A

C--> 29 (B) FILING DATE: 24-Aug-2001

30 (C) CLASSIFICATION: 435

32 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: 08/760,745

34 (B) FILING DATE: 1996-12-07

36 (viii) ATTORNEY/AGENT INFORMATION:

37 (A) NAME: Billings, Lucy J.

38 (B) REGISTRATION NUMBER: 36,749

39 (C) REFERENCE/DOCKET NUMBER: PF-0169 US

41 (ix) TELECOMMUNICATION INFORMATION:

42 (A) TELEPHONE: 415-855-0555

43 (B) TELEFAX: 415-845-4166

44 (C) TELEX:

47 (2) INFORMATION FOR SEQ ID NO: 1:

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 235 amino acids

51 (B) TYPE: amino acid

52 (C) STRANDEDNESS: single

53 (D) TOPOLOGY: linear

55 (vii) IMMEDIATE SOURCE:

56 (A) LIBRARY: lungast01

57 (B) CLONE: 876242

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 Met Pro His Ala Phe Lys Pro Gly Asp Leu Val Phe Ala Lys Met Lys

62 1 5 10 15

63 Gly Tyr Pro His Trp Pro Ala Arg Ile Asp Asp Ile Ala Asp Gly Ala

ENTERED

RECEIVED

DEC 1 8 2001

TECH CENTER 1600/2900

RAW SEQUENCE LISTING

DATE: 11/29/2001

PATENT APPLICATION: US/09/938,885A

TIME: 11:02:26

Input Set : N:\Crf3\RULE60\09938885A.RAW

Output Set: N:\CRF3\11292001\I938885A.raw

```

64          20          25          30
65 Val Lys Pro Pro Pro Asn Lys Tyr Pro Ile Phe Phe Phe Gly Thr His
66          35          40          45
67 Glu Thr Ala Phe Leu Gly Pro Lys Asp Leu Phe Pro Tyr Glu Glu Ser
68          50          55          60
69 Lys Glu Lys Phe Gly Lys Pro Asn Lys Arg Lys Gly Phe Ser Glu Gly
70          65          70          75          80
71 Leu Trp Glu Ile Glu Asn Asn Pro Thr Val Lys Ala Ser Gly Tyr Gln
72          85          90          95
73 Ser Ser Gln Lys Lys Ser Cys Val Glu Glu Pro Glu Pro Glu Pro Glu
74          100         105         110
75 Ala Ala Glu Gly Asp Gly Asp Lys Lys Gly Asn Ala Glu Gly Ser Ser
76          115         120         125
77 Asp Glu Glu Gly Lys Leu Val Ile Asp Glu Pro Ala Lys Glu Lys Asn
78          130         135         140
79 Glu Lys Gly Ala Leu Lys Arg Arg Ala Gly Asp Leu Leu Glu Asp Ser
80          145         150         155         160
81 Pro Lys Arg Pro Lys Glu Ala Glu Asn Pro Glu Gly Glu Glu Lys Glu
82          165         170         175
83 Ala Ala Thr Leu Glu Val Glu Arg Pro Leu Pro Met Glu Val Glu Lys
84          180         185         190
W--> 85 Asn Ser Thr Pro Ser Glu Pro Gly Ser Gly Arg Gly Pro Pro Xaa Xaa
86          195         200         205
W--> 87 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Glu Glu Ala Thr Lys Glu Asp Ala
88          210         215         220
89 Glu Ala Pro Gly Ile Lys Ser His Glu Ser Leu
90          225         230         235

```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 869 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: lungast01

(B) CLONE: 876242

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

106 CCGCTGCAGC CGCTTTCTGC GGCCTGGGCC TCTCGCCGTC AGCATGCCAC ACGCCTTCAA 60
107 GCCCGGGGAC TTGGTGTTTC CTAAGATGAA GGGCTACCTT CACTGGCCTG CCAGGATCGA 120
108 CGACATCGCG GATGGCGCCG TGAAGCCCCC ACCCAACAAG TACCCCATCT TTTTCTTTGG 180
109 CACACACGAA ACAGCCTTCC TGGGCCCCAA AGACCTCTTC CCTTACGAGG AATCCAAGGA 240
110 GAAGTTTGGC AAGCCCAACA AGAGGAAAGG GTTCAGCGAG GGGCTGTGGG AGATCGAGAA 300
111 CAACCTACT GTCAAGGCTT CCGGCTATCA GTCTCCAG AAAAAAGAGCT GTGTGGAAGA 360
112 GCCTGAACCA GAGCCGAAG CTGCAGAGGG TGACGGTGAT AAGAAGGGGA ATGCAGAGGG 420
113 CAGCAGCGAC GAGGAAGGGA AGCTGGTCAT TGATGAGCCA GCCAAGGAGA AGAACGAGAA 480
114 AGGAGCGTTG AAGAGGAGAG CAGGGGACTT GCTGGAGGAC TCTCCTAAAC GTCCCAAGGA 540
115 GGCAGAAAAC CCTGAAGGAG AGGAGAAGGA GGCAGCCACC TTGGAGGTTG AGAGGCCCTT 600
116 TCCTATGGAG GTGGAAGA ATAGCACCCC CTCTGAGCCC GGCTCTGGCC GGGGGCCTCC 660
117 CCNNNNNNNN NNNNNNNNNN NNNNNNNNNA GGAAGAGGCT ACCAAGGAAG ATGCTGAGGC 720

```

RAW SEQUENCE LISTING

DATE: 11/29/2001

PATENT APPLICATION: US/09/938,885A

TIME: 11:02:26

Input Set : N:\Crf3\RULE60\09938885A.RAW

Output Set: N:\CRF3\11292001\I938885A.raw

```

118 CCCAGGCATC AAGAGTCATG AGAGCCTGTA GCCACCAATG TTTCAAGAGG AGCCCCCACC 780
119 CTGTTCTCTGC TGCTGTCTGG GTGCTACTGG GGAAACTGGC CATGGGCTGC AAACCTGGGNA 840
120 CCCCTTTTCC ANCNCAANCT GNTNTTCTT 869

```

122 (2) INFORMATION FOR SEQ ID NO: 3:

124 (i) SEQUENCE CHARACTERISTICS:

125 (A) LENGTH: 240 amino acids

126 (B) TYPE: amino acid

127 (C) STRANDEDNESS: single

128 (D) TOPOLOGY: linear

130 (vii) IMMEDIATE SOURCE:

131 (A) LIBRARY: GenBank

132 (B) CLONE: 598956

134 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

136 Met Ser Arg Ser Asn Arg Gln Lys Glu Tyr Lys Cys Gly Asp Leu Val
137 1 5 10 15
138 Phe Ala Lys Met Lys Gly Tyr Pro His Trp Pro Ala Arg Ile Asp Glu
139 20 25 30
140 Met Pro Glu Ala Ala Val Lys Ser Thr Ala Asn Lys Tyr Gln Val Phe
141 35 40 45
142 Phe Phe Gly Thr His Glu Thr Ala Phe Leu Gly Pro Lys Asp Leu Phe
143 50 55 60
144 Pro Tyr Glu Glu Ser Lys Glu Lys Phe Gly Lys Pro Asn Lys Arg Lys
145 65 70 75 80
146 Gly Phe Ser Glu Gly Leu Trp Glu Ile Glu Asn Asn Pro Thr Val Lys
147 85 90 95
148 Ala Ser Gly Tyr Gln Ser Ser Gln Lys Lys Ser Cys Val Glu Glu Pro
149 100 105 110
150 Glu Pro Glu Pro Glu Ala Ala Glu Gly Asp Gly Asp Lys Lys Gly Asn
151 115 120 125
152 Ala Glu Gly Ser Ser Asp Glu Glu Gly Lys Leu Val Ile Asp Glu Pro
153 130 135 140
154 Ala Lys Glu Lys Asn Glu Lys Gly Ala Leu Lys Arg Arg Ala Gly Asp
155 145 150 155 160
156 Leu Leu Glu Asp Ser Pro Lys Arg Pro Lys Glu Ala Glu Asn Pro Glu
157 165 170 175
158 Gly Glu Glu Lys Glu Ala Ala Thr Leu Glu Val Glu Arg Pro Leu Pro
159 180 185 190
160 Met Glu Val Glu Lys Asn Ser Thr Pro Ser Glu Pro Gly Ser Gly Arg
161 195 200 205
162 Gly Pro Pro Gln Glu Glu Glu Glu Glu Asp Glu Glu Glu Glu Ala
163 210 215 220
164 Thr Lys Glu Asp Ala Glu Ala Pro Gly Ile Arg Asp His Glu Ser Leu
165 225 230 235 240

```

167 (2) INFORMATION FOR SEQ ID NO: 4:

169 (i) SEQUENCE CHARACTERISTICS:

170 (A) LENGTH: 2376 base pairs

171 (B) TYPE: nucleic acid

172 (C) STRANDEDNESS: single

173 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING

DATE: 11/29/2001

PATENT APPLICATION: US/09/938,885A

TIME: 11:02:26

Input Set : N:\Crf3\RULE60\09938885A.RAW

Output Set: N:\CRF3\11292001\I938885A.raw

175 (vii) IMMEDIATE SOURCE:

176 (A) LIBRARY: GenBank

177 (B) CLONE: 598956

179 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

181 GAGGAGGAGT GGGGACCGGG CGGGGGGTGG AGGAAGAGGC CTCGCGCAGA GGAGGGAGCA      60
182 ATTGAATTC AAACACAAAC AACTCGACGA GCGCGCACCC ACCGCGCCGG AGCCTTGCCC      120
183 CGATCCGCGC CCGCCCCGTC CGTGCGGCGC GCGGGCGGAG ACGCCGTGGC CGCGCCGGAG      180
184 CTCGGGCCGG GGGCCACCAT CGAGGCGGGG GCCGCGCGAG GGCCGGAGCG GAGCGGCGCC      240
185 GCCACCGCCG CACGCGCAAA CTTGGGCTCG CGCTTCCCGG CCGGCGCGG AGCCCGGGGC      300
186 GCCCGGAGCC CCGCCATGTC GCGATCCAAC CGGCAGAAGG AGTACAAATG CGGGGACCTG      360
187 GTGTTCGCCA AGATGAAGGG CTACCCACAC TGGCCGGCCC GGATTGACGA GATGCCTGAG      420
188 GCTGCGGTGA AATCAACAGC CAACAAATAC CAAGTCTTTT TTTTCGGGAC CCACGAGACG      480
189 GCATTCTTGG GCCCCAAAGA CCTCTTCCCT TACGAGGAAT CCAAGGAGAA GTTTGGCAAG      540
190 CCAACAAGA GGAAGGGT CAGCGAGGGG CTGTGGGAGA TCGAGAACAA CCCTACTGTC      600
191 AAGGCTTCCG GCTATCAGTC CTCCAGAAA AAGAGCTGTG TGGAAGAGCC TGAACCAGAG      660
192 CCCGAAGCTG CAGAGGGTGA CGGTGATAAG AAGGGGAATG CAGAGGGCAG CAGCGACGAG      720
193 GAAGGGAAGC TGGTCATTGA TGAGCCAGCC AAGGAGAAGA ACGAGAAAGG AGCGTTGAAG      780
194 AGGAGAGCAG GGGACTTGCT GGAGGACTCT CCTAACGTC CCAAGGAGGC AGAAAAACCT      840
195 GAAGGAGAGG AGAAGGAGGC AGCCACCTTG GAGGTTGAGA GGCCCCCTCC TATGGAGGTG      900
196 GAAAAAGAATA GCACCCCTC TGAGCCCGGC TCTGGCCGGG GGCCCTCCCA AGAGGAAGAA      960
197 GAAGAGGAGG ATGAAGAGGA AGAGGCTACC AAGGAAGATG CTGAGGCCCC AGGCATCAGA     1020
198 GATCATGAGA GCCTGTAGCC ACCAATGTTT CAAGAGGAGC CCCACCCCTG TTCCTGCTGC     1080
199 TGTCTGGGTG CTA CTG GGGG AACTGGCCAT GGCTGCAAAA CTGGGAACCC CTTTCCACC     1140
200 CCAACCTGCT CTCCTCTTCT ACTCACTTTT CCCACTCCAA GCCCAGCCCA TGGAGATTGA     1200
201 CCTGGATTGG GCAGGCCACC TGGCTCTCAC CTCTAGGTCC CCATACTCCT ATGATCTGAG     1260
202 TCAGAGCCAT GTCTTCTCCC TGGATAGT TGAGGCACT GTGTTCCTTC CGCTTGGAGC     1320
203 TATTTTCCAG GCTTCTGCTG GGGCCTGGGA CAACTGCTCC CACCTCCTGA CACCTTCTC     1380
204 CCACTCTCCT AGGCATTCTG GACCTCTGGG TTGGGATCAG GGGTAGGAAT GGAAGGATGG     1440
205 AGCATCAACA GCAGGGTGGG CTTGTGGGGC CTGGGAGGGG CAATCCTCAA ATGCGGGGTG     1500
206 GGGGAGCAC AGGAGGGCGG CCTCCTTCTG AGCTCCTGTC CCCTGCTACA CCTATTATCC     1560
207 CAGCTGCCTA GATTGAGGA AAGTGGGACA GCTTGAGGGG GAGGGGCTCC TTTCCATAAA     1620
208 TCCTTGATGA TTGACAACAC CCATTTTTCC TTTTGCCGAC CCAAGAGTT TTGGGAGTTG     1680
209 TAGTTAATCA TCAAGAGAAT TTGGGGCTTC CAAGTTGTTC GGGCCAAGGA CCTGAGACCT     1740
210 GAAGGGTTGA CTTTACCCAT TTGGGTGGGA GTGTGAGCA TCTGTCCCCC TTTAGATCTC     1800
211 TGAAGCCACA AATAGGATGC TTGGGAAGAC TCCTAGCTGT CCTTTTTCCT CTCCACACAG     1860
212 TGCTCAAGGC CAGCTTATAG TCATATATAT CACCCAGACA TAAAGGAAAA GACACATTTT     1920
213 TTAGGAAATG TTTTAAATAA AAGAAAATTA CAAAAAATAA TTTTAAAGAC CCCTAACCTT     1980
214 TTGTGTGCTC TCCATTCTGC TCCTTCCCCA TCGTTGCCCC CATTTCTGAG GTGCACTGGG     2040
215 AGGCTCCCTT TCTATTTGGG GCTTGATGAC TTTCTTTTTG TAGCTGGGGC TTTGATGTTT     2100
216 CTTCCAGTGT CATTTCTCAT CCACATACCC TGACCTGGCC CCCTCAGTGT TGTCACCAGA     2160
217 TCTGATTGTG AACCACTGA GAGGACAGAG AGAAATAAGT GCCCTCTCCC ACCCTCTTCC     2220
218 TACTGGTCTC TCTATGCCTC TCTACAGTCT CGTCTCTTTT ACCCTGGCCC CTCTCCCTTG     2280
219 GGCTCTGATG AAAAATTGCT GACTGTAGCT TTGGAAGTTT AGCTCTGAGA ACCGTAGATG     2340
220 ATTTACAGTT TAGGAAAATA AAACCCGTTG ATTACT                                2376

```

222 (2) INFORMATION FOR SEQ ID NO: 5:

224 (i) SEQUENCE CHARACTERISTICS:

225 (A) LENGTH: 237 amino acids

226 (B) TYPE: amino acid

227 (C) STRANDEDNESS: single

RAW SEQUENCE LISTING

DATE: 11/29/2001

PATENT APPLICATION: US/09/938,885A

TIME: 11:02:26

Input Set : N:\Crf3\RULE60\09938885A.RAW

Output Set: N:\CRF3\11292001\I938885A.raw

```

228      (D) TOPOLOGY: linear
230      (vii) IMMEDIATE SOURCE:
231      (A) LIBRARY: GenBank
232      (B) CLONE: 945419
234      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
236 Met Ser Arg Ser Asn Arg Gln Lys Glu Tyr Lys Cys Gly Asp Leu Val
237   1          5          10          15
238 Phe Ala Lys Met Lys Gly Tyr Pro His Trp Pro Ala Arg Ile Asp Glu
239          20          25          30
240 Met Pro Glu Ala Ala Val Lys Ser Thr Ala Asn Lys Tyr Gln Val Phe
241          35          40          45
242 Phe Phe Gly Thr His Glu Thr Ala Phe Leu Gly Pro Lys Asp Leu Phe
243          50          55          60
244 Pro Tyr Glu Glu Ser Lys Glu Lys Phe Gly Lys Pro Asn Lys Arg Lys
245          65          70          75          80
246 Gly Phe Ser Glu Gly Leu Trp Glu Ile Glu Asn Asn Pro Thr Val Lys
247          85          90          95
248 Ala Ser Gly Tyr Gln Ser Ser Gln Lys Lys Ser Cys Ala Ala Glu Pro
249          100         105         110
250 Glu Val Glu Pro Glu Ala His Glu Gly Asp Gly Asp Lys Lys Gly Ser
251          115         120         125
252 Ala Glu Gly Ser Ser Asp Glu Glu Gly Lys Leu Val Ile Asp Glu Pro
253          130         135         140
254 Ala Lys Glu Lys Asn Glu Lys Gly Thr Leu Lys Arg Arg Ala Gly Asp
255          145         150         155         160
256 Val Leu Glu Asp Ser Pro Lys Arg Pro Lys Glu Ser Gly Asp His Glu
257          165         170         175
258 Glu Glu Asp Lys Glu Ile Ala Ala Leu Glu Gly Glu Arg His Leu Pro
259          180         185         190
260 Val Glu Val Glu Lys Asn Ser Thr Pro Ser Glu Pro Asp Ser Gly Gln
261          195         200         205
262 Gly Pro Pro Ala Glu Glu Glu Glu Gly Glu Glu Glu Ala Ala Lys Glu
263          210         215         220
264 Glu Ala Glu Ala Pro Gly Val Arg Asp His Glu Ser Leu
265          225         230         235

```

VERIFICATION SUMMARY

DATE: 11/29/2001

PATENT APPLICATION: US/09/938,885A

TIME: 11:02:27

Input Set : N:\Crf3\RULE60\09938885A.RAW

Output Set: N:\CRF3\11292001\I938885A.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:85 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1

L:87 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1